

Conference Abstract

Angling for data: making biodiversity metadata more FAIR

Joakim Philipson ‡

‡ Stockholm University, Stockholm, Sweden

Corresponding author: Joakim Philipson (jomtov@yahoo.com)

Received: 14 Aug 2017 | Published: 14 Aug 2017

Citation: Philipson J (2017) Angling for data: making biodiversity metadata more FAIR. Proceedings of TDWG 1: e20267. <https://doi.org/10.3897/tdwgproceedings.1.20267>

Abstract

The **FAIR** guiding principles, first launched in 2014, for making research data more **Findable**, **Accesible**, **Interoperable** and **Re-usable**, have not yet been widely implemented for biodiversity data. Partly this may be due to the FAIR principles by themselves not yet being fully operational and easy to interpret. There is work in progress to remedy this by different task groups, and different attempts have already been made. In this paper I will give some concrete tips aimed at implementing the FAIR principles for biodiversity research data, focusing on the metadata, in order to enhance the quality of data by making them more findable, accessible, interoperable and reusable. Among the steps that could be taken to make biodiversity database records more **findable** and **accessible** is for example to add schema.org markup to the html sourcecode of corresponding web pages, as has been successfully employed in the *Uniprot* database. Recently biocaddie.org has mapped the metadata format DATS, *Data Tag Suite*, to schema.org and there is also the ongoing adaptation effort of bioschemas.org. In addition, there is the highly commendable work done by former biosharing.org, which now has become the more general fairsharing.org and which aims to enhance findability, promote the adoption of metadata standards by policy makers and interlink metadata standards among themselves and with repositories (Sansone 2017).

Further, to make biodiversity records more **interoperable** and **reusable**, it is essential to provide metadata export to a selection of general standards and formats. In doing this, promises should be kept, meaning that exported metadata records should also **validate**

against the schemas for the chosen format standard. By validating against schemas of both preferred metadata standard and export formats, biodiversity data records also stand a better chance of achieving what has been defined by GBIF and Vertnet as *Fitness-for-use*, encompassing e.g. *accessibility, content, completeness, dataset-level or record level, error correction* etc. (Russell 2011). That is, of course, provided the relevant metadata standards have validation schemas or online tools such as the [Darwin Core Archive/EML validator](#) that are sufficiently precise to check for these properties. If not, there is always the possibility of creating tailor-made validation schemas serving the data quality needs of a specialized biodiversity data repository, e.g. using Schematron or JSON schema.

Keywords

FAIR principles, metadata, validation

Presenting author

Joakim Philipson (Stockholm University Library)

Presented at

TDWG 2017 Annual Conference, Symposium: Biodiversity Data Quality – concepts, methods and tools

Hosting institution

Stockholm University Library

References

- Russell L (2011) Brief introduction to metadata in a data quality context. <https://vimeo.com/album/1904479/video/40447148>. Accessed on: 2017-7-19.
- Sansone S (2017) From BioSharing to FAIRsharing - mapping the standards landscape. <https://www.slideshare.net/SusannaSansone/from-biosharing-to-fairsharing-mapping-the-standards-landscape>. Accessed on: 2017-7-18.